A Nonparametric Test For Homogeneity Of Variances: Application To GPAs Of Students Across Academic Majors

Saad T. Bakir, Alabama State University, USA

ABSTRACT

We propose a nonparametric (or distribution-free) procedure for testing the equality of several population variances (or scale parameters). The proposed test is a modification of Bakir's (1989, Commun. Statist., Simul-Comp., 18, 757-775) analysis of means by ranks (ANOMR) procedure for testing the equality of several population means. A proof is given to establish the distribution-free property of the modified procedure. The proposed procedure is then applied to test whether or not the variability in the grade point averages (GPAs) of students differs across five business academic majors. We collect the GPAs (observations) of a random sample of students from each major under study. The absolute deviations of the observations from the overall median of the combined sample are then calculated and ranked from least to largest. The average ranks and two decision lines are then plotted on a graph paper to detect not only the existence of significant differences among variances, but also to pinpoint which variances are causing those differences.

Keywords: ANOM, ANOMR, Distribution-free, Multi-sample scale-problem

1. INTRODUCTION

The analysis of means (ANOM) procedure was originally proposed by Ott (1967) as alternative to the analysis of variance (ANOVA) test for the equality of means of several normal populations. A comprehensive account of several ANOM-type procedures can be found in Nelson et al. (2005). Bakir (1989, 1994) developed nonparametric (or distribution-free) versions of ANOM to test the equality of several means in the settings of a completely randomized design and a randomized complete block design. Based on Bakir's (1989) ANOMR version, we develop in this paper a nonparametric procedure for testing the null hypothesis of equal population variances (or scale parameters). In a truly distribution-free test procedure, the distribution (called the null distribution) of the test statistic under the null hypothesis, must not depend on the functional form of the underlying parent distribution of the observations. When testing the equality of only two variances, our proposed procedure is equivalent to Flinger and Killeen (1976) distribution-free two-sample test for scale.

The most general setting for the problem of testing the equality of several variances (the so called multi-sample scale problem, or homogeneity of variances problem) can be outlined as follows:
For $1 \le i \le I$, let $\left(X_{i1}, X_{i2}, \cdots, X_{in_i}\right)$ be mutually independent random samples drawn from populations with cumulative distribution functions (CDFs) $F_1, F_2, \cdots, F_I$, where,

$$F_i\left(x\right) = F\left(\frac{x-\mu_i}{\sigma_i}\right), 1 \le i \le I. \qquad (1)$$

The function F is an unknown absolutely continuous CDF that involves a location parameter $-\infty < \mu_i < \infty$ and a scale parameter $\sigma_i > 0$. When $I = 2$, we have the so called two-sample scale problem. The most general null and alternative hypotheses to be tested in the multi-sample scale problem, respectively, are:

$$H_0 : \sigma_1 = \sigma_2 = \cdots = \sigma_I \tag{2}$$

$$H_1 : \sigma_i \neq \sigma_{i'} \text{ for some } i \neq i'. \tag{3}$$

To be truly distribution-free, however, most proposed nonparametric procedures require the assumption that the populations have known location parameters or, have the same (unknown) location parameter. The hypotheses of interest would then become:

$$H_0 : \sigma_1 = \sigma_2 = \cdots = \sigma_I ,\ \mu_i = \mu , \tag{4}$$

$$H_1 : \sigma_i \neq \sigma_{i'} ,\ \mu_i = \mu . \tag{5}$$

The constant μ represents the common (unknown) location parameter, usually representing a central value (the median, or the mean if it exists) of the populations. Deshpande and Kusum (1984) and Kusum (1985) considered the two-sample scale problem when the common unknown location parameter is a general quantile of the populations. If the population centers are unknown and unequal, then Moses (1963) and Blair and Thompson (1992) rank-like procedures are among the very few truly distribution-free tests for the two-sample scale problem. For a literature review on the two- and multi-sample scale problems, see Duran (1976), Daniel (1979), Conover et al. (1981) and Shetty et al. (2004). To better understand the proposed procedure, we first need to summarize (in Section 2) Bakir's (1989) ANOMR procedure for testing the equality of several population means.

2. ANOMR FOR TESTING THE EQUALITY OF MEANS

With reference to the setup in Eq. (1), ANOMR was proposed by Bakir (1989) as a distribution-free procedure to test the following null hypothesis of equality of several population central parameters (means, or medians):

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_I = \mu, \sigma_i = \sigma , \tag{6}$$

against the alternative hypothesis:

$$H_a : \mu_i \neq \mu_{i'} \ \mathit{for}\ i \neq i', \sigma_i = \sigma . \tag{7}$$

In the ANOMR procedure, the original $X_{ij}, \left(1 \le i \le I, 1 \le j \le n_i\right)$, observations themselves are ranked from least to largest in the combined sample. For clarity, we use "*X*" to label the various quantities in Bakir's (1989) ANOMR procedure. The ranks and average ranks are given by:

$$R_{ij}^X = \text{rank of } \mathrm{X}_{ij} \text{ in the combined sample of X's}, \left(1 \le i \le I, 1 \le j \le n_i\right) \tag{8}$$

$$\bar{R}_i^X = \sum_{j=1}^{n_i} \left(R_{ij}^X / n_i \right), 1 \le i \le I \tag{9}$$

$$\bar{\bar{R}}^X = \sum_{i=1}^{I}\sum_{j=1}^{n_i} R_{ij}^X / N = (N+1)/2, \tag{10}$$

where $N = \sum_{i=1}^{I} n_i$.

The ANOMR test rejects H_0 in Eq. (6) if for any *i*, we get

$$\left|\bar{R}_i^X - \bar{\bar{R}}^X\right| \geq C^X \tag{11}$$

Equivalently, the ANOMR test rejects H_0 if

$$\max_{1\leq i\leq I}\left|R_i^X - \bar{\bar{R}}^X\right| \geq C^X . \tag{12}$$

When H_0 in Eq. (6) is true, the quantities $R_{ij}^X, (1 \leq i \leq I, 1 \leq j \leq n_i)$, being ranks of independently and identically distributed (*iid*) random variables, the X_{ij} 's, have a discrete uniform distribution irrespective of the common underlying parent distribution $F\left(\frac{x-\mu}{\sigma}\right)$ of the observations. Therefore, ANOMR is a truly a distribution-free test because any test based on the ranks of *iid* random variables is distribution-free; see Randles and Wolfe (1979, corollary 2.3.6 pp 39). For values of $I = 3,4$ and certain values of the n_i 's, Bakir (1989) calculated the exact (and large sample approximate) values of $C^X \equiv C^X(\alpha; I, n_1, n_2, \cdots, n_I)$ such that

$$\Pr\left[\max_{1\leq i\leq I}\left|\bar{R}_i^X - \bar{\bar{R}}^X\right| \geq C^X\right] = \alpha . \tag{13}$$

Further, Bakir (1989) developed large sample (asymptotic) and other type of approximations for the critical values of the ANOMR procedure. The asymptotic values of C^X are calculated via the formula

$$C^X = [(I-1)(N+1)/12]^{1/2}\,\omega(\alpha; I), \tag{14}$$

where, $\omega(\alpha; I)$ is to be read from Table IV of Bakir (1989) depending on the level of significance, α, and the number, *I*, of populations being compared.

3. A NONPARAMETRIC ANOMR-TYPE TEST FOR HOMOGENEITY OF VARIANCES

In this section we develop a nonparametric ANOM-type procedure for testing the equality of several variances (the multi-sample scale problem). We need to assume that the populations have a common (unknown) center (median or mean). The procedure is designed to test the null and alternative hypotheses in Eqs. (4) and (5).

Let the observations $X_{ij}, (1 \leq i \leq I, 1 \leq j \leq n_i)$, be as defined in Section 1 and denote the mean and median of the combined sample, respectively, by $\bar{\bar{X}}$ and $\tilde{\tilde{X}}$. The proposed procedure hinges on replacing the observations, X_{ij} , by their absolute deviations from the median of the combined sample. Those absolute deviations $\left|X_{ij}-\tilde{\tilde{X}}\right|,\left(1\le i\le I,1\le j\le n_i\right)$, are then ranked from least to largest in the combined sample. ANOMR, as outlined in Section 2, is then applied to those absolute ranks.

Although both of the transformations $\left|X_{ij}-\tilde{\tilde{X}}\right|$ and $\left(X_{ij}-\bar{\bar{X}}\right)^2$ produce the same ranks in the combined sample, we use the $\left|X_{ij}-\tilde{\tilde{X}}\right|$ transformations because they are easier to calculate. In Section 3.1 we develop the procedure and in Section 3.2 we establish its distribution-free property.

3.1 Development Of The Anomr-Type Homogeneity Of Variances Procedure

Define the transformations:

$$U_{ij}=\left|X_{ij}-\tilde{\tilde{X}}\right|,\left(1\le i\le I,1\le j\le n_i\right). \tag{15}$$

Define the following ranks and average ranks:

$$R_{ij}^U=\text{rank of } U_{ij} \text{ in the combined sample of } U's,\left(1\le i\le I,1\le j\le n_i\right) \tag{16}$$

$$\bar{R}_i^U=\sum_{j=1}^{n_i}\left(R_{ij}^U/n_i\right)\qquad 1\le i\le I \tag{17}$$

$$\bar{\bar{R}}^U=\sum_{i=1}^{I}\sum_{j=1}^{n_i}R_{ij}^U/N=\left(N+1\right)/2 \tag{18}$$

The proposed test rejects H_0 in Eq. (4) if for any *i*, we get

$$\left|\bar{R}_i^U-\bar{\bar{R}}^U\right|\ge C^U \tag{19}$$

Equivalently, the proposed test rejects H_0 if

$$\max_{1\le i\le I}\left|R_i^U-\bar{\bar{R}}^U\right|\ge C^U \tag{20}$$

The critical values C^U do not require special tables; they are identical to the exact and the approximate values C^X , discussed in Section 2. When *I =2*, the proposed test procedure is equivalent to Flinger and Killeen (1976) distribution-free two-sample test for scale.

The proposed procedure can be carried out graphically by plotting the points $\left(i,\bar{R}_i^u\right)$ on a graph and marking the following upper decision line (UDL), lower decision line (LDL), and center line (CL):

$$UDL = \overline{\overline{R}}^{u} + C^{u}, \quad LDL = \overline{\overline{R}}^{u} - C^{u}, \text{ and } CL = \overline{\overline{R}}^{u} \tag{21}$$

The null hypothesis of equal variances is rejected if any of the plotted points falls outside the upper and lower decision lines; otherwise the null hypothesis is not rejected.

3.2 A Proof Of The Distribution-Free Property Of The Proposed Procedure

In this section we demonstrate that our proposed test is based on ranks of exchangeable random variables, the U_{ij}'s, and, hence, is distribution-free. For a definition of exchangeable random variables, see Randles and Wolfe (1979, Definition 1.3.6, pp 15).

To prove exchangeability, we express the U_{ij}'s in Eq. (15) as

$$U_{ij} = h\left(X_{ij}; g\left(X_{ij}\right)\right), \left(1 \le i \le I, 1 \le j \le n_i\right), \tag{22}$$

where $g\left(X_{ij}\right) = \tilde{\tilde{X}}$ and $h\left(t_1, t_2\right) = \left|t_1 - t_2\right|$. (23)

Since the X_{ij}'s are *iid* and $g\left(X_{ij}\right)$ is symmetric in its arguments, then the U_{ij}'s are exchangeable random variables and any test based on their ranks would be distribution-free; see Randles and Wolfe (1979, Theorem 11.2.3 and Corollary 11.2.5 pp 357). Therefore, the proposed test is distribution-free.

It is to be noted that Wludyka and Nelson (1999) proposed tests for homogeneity of variances as viable nonparametric procedures. However, their procedures are not truly distribution-free because they are based on ranks of non-independent and non-exchangeable random variables.

4. AN APPLICATION TO TEST VARIABILITY IN THE GPAs OF STUDENTS

We now apply the proposed procedure, as developed in Section 3, to test the null hypothesis of equal variances of the grade point averages (GPAs) across five business academic majors: accounting (ACT), computer information systems (CIS), finance (FIN), management (MGT), and marketing (MKT). In symbols, we want to test the null and alternative hypotheses in Eq. (4) and Eq. (5).

At the end of summer 2009, a random sample of 10 students was selected from each one of the five business majors of our college. Table 1 shows the cumulative GPA of the selected students together with the overall median (Med) of the combined sample. Table 2 shows the absolute difference (deviation) of each GPA from the overall median. The absolute deviations are then ranked from least to largest in the combined sample and the average ranks are calculated as shown in Table 3.

Table 1. GPAs of Random Samples of Students in Five Business Majors as of Summer2009

ACT	CIS	FIN	MGT	MKT
4.000	3.561	3.063	2.311	2.500
2.633	2.729	2.376	3.425	2.367
2.253	2.694	3.406	2.541	2.517
2.063	3.066	3.667	2.744	2.798
3.741	3.914	2.509	2.348	3.509
2.463	2.414	2.286	2.821	2.371
3.103	2.048	3.800	2.667	3.080
2.576	2.368	2.442	2.566	2.488
2.962	2.828	3.278	2.456	2.630
2.289	2.464	2.698	3.126	2.262
Overall median of the combined sample Med = 2.632				

Table 2. Absolute Deviations of the GPAs from the Median (Med)

ACT	CIS	FIN	MGT	MKT
1.369	0.930	0.432	0.321	0.132
0.002	0.098	0.256	0.794	0.265
0.379	0.063	0.775	0.091	0.115
0.569	0.435	1.036	0.113	0.167
1.110	1.283	0.123	0.284	0.878
0.169	0.218	0.346	0.190	0.261
0.472	0.584	1.169	0.035	0.449
0.055	0.264	0.190	0.066	0.144
0.331	0.197	0.647	0.176	0.002
0.343	0.168	0.067	0.495	0.370

Table 3. Ranks of the Absolute Deviations in the Combined Sample

	ACT	CIS	FIN	MGT	MKT
	50	45	34	28	13
	1	9	23	43	26
	33	5	42	8	11
	39	35	46	10	15
	47	49	12	27	44
	17	22	31	20	24
	37	40	48	3	36
	4	25	19	6	14
	29	21	41	18	1
	30	16	7	38	32
$\bar{R}^u$	28.7	26.7	30.3	20.1	21.6
$\bar{\bar{R}}^u$	25.48				

Reading Table IV of Bakir (1989) with significance level $\alpha = 0.05$ and number of populations *I=5*, we find $\omega(\alpha; I) = 2.56$. The asymptotic critical value for our procedure then becomes

$$C^U = [(I-1)(N+1)/12]^{1/2}\,\omega(\alpha; I) = [(4)(51)/12]^{1/2}(2.56) = 10.56.$$

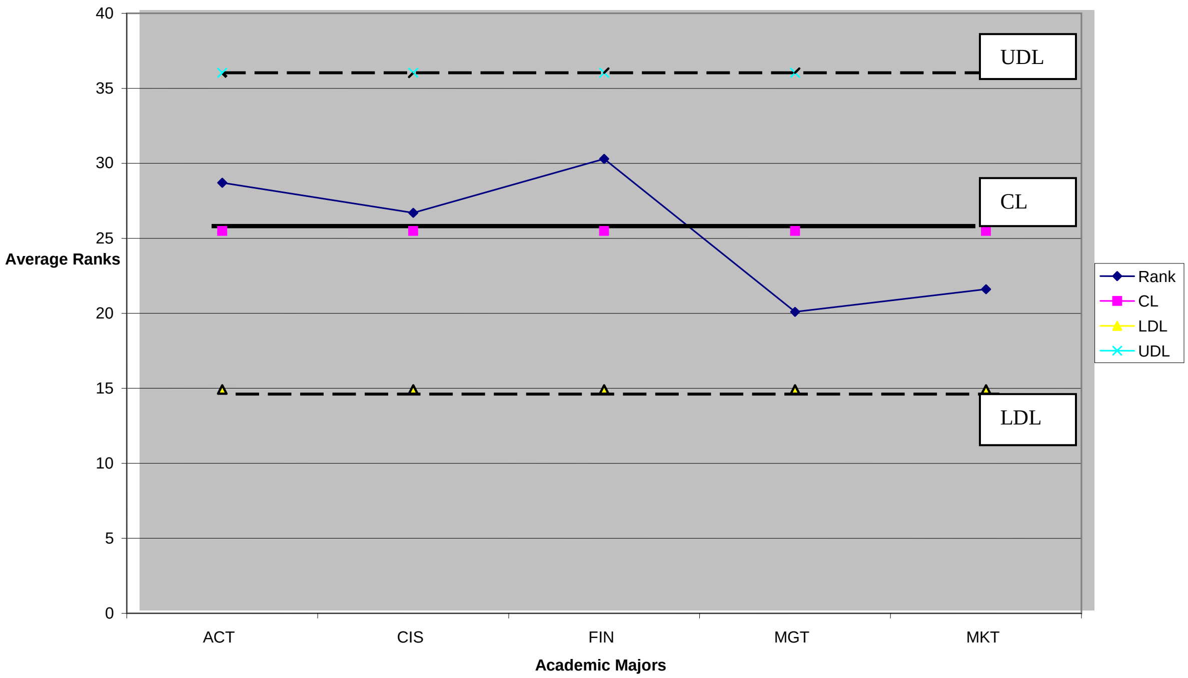
From Table 3, we get the value of the center line $\bar{\bar{R}}^U = 25.48$, and calculate the following upper and lower decision lines:

$$UDL = \bar{\bar{R}}^u + C^u = 25.48 + 10.56 = 36.04,$$

$$LDL = \bar{\bar{R}}^u - C^u = 25.48 - 10.56 = 14.92.$$

Plotting the values of the average ranks of the five majors (see Table 3) and marking the center and decision lines result in the chart of Figure 1.

Figure 1. A Nonparametric ANOMR Chart for Variances



Since all the plotted points (the average ranks of the five academic majors) on the chart of Figure 1 fall within the upper and lower decision lines, the null hypothesis of equal variances is not rejected. We conclude that there is no significant difference in the variability of GPAs across academic majors

5. CONCLUSION

In this paper we reviewed the analysis of means by ranks (ANOMR) procedure proposed by Bakir (1989) for testing the equality of several population means. The ANOMR procedure is then modified to test the equality of several population variances (also known as the homogeneity of variances, or the multi-sample scale problem). The modified ANOMR procedure is proved to be a distribution-free (or nonparametric) test. The proposed procedure can be carried out graphically to visualize the positions of variances being compared. As an application of the proposed procedure, a study was conducted to test whether or not the variability in the grade point averages (GPAs) of students differ across five business academic majors. The analysis revealed no significant difference in the variability of GPAs across the academic majors

REFERENCES

1. Ansari, A. R. and Bradley, R. A. (1960). Rank sum tests for dispersion. Ann. *Math. Statist*. 31, 1174-1189.
2. Bakir, S. T. (1989). Analysis of means using ranks. *Communications in Statistics-Simulation and Computation*, 18, 757-775.
3. Bakir, S. T. (1994). Analysis of means using ranks for the randomized complete block design. *Communications in Statistics-Simulation and Computation*, 23(2), 547-568.
4. Blair, R. C. and Thompson, G. L. (1992). A distribution-free rank-like test for scale with unequal population locations. *Communications in Statistics-Simulation and Computation*, 21(2), 353-371.
5. Conover, W. J., Johnson, M. E., and Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances with applications to the outer continental shelf bidding data. *Technometrics*, 23, 351-361.
6. Daniel, W. W. (1979). On nonparametric and robust tests for dispersion: a selected bibliography. Public administration series-bibliography, P-382, Vance Bibliographies, Monticello, III. (5.3).

7. Dishbande, J V. and Kusum, K. (1984). A test for the nonparametric two sample scale problem. *Australian Journal of Statistics*, 26(1), 16-24.
8. Duran, B. S. (1976). A survey of Nonparametric tests for scale. *Communications in Statistics-Theory and Methods*. A 5, 1287-1312.
9. Fligner, M. A. and Killeen, T. J (1976). Distribution-free two-sample tests for scale. *J. of the American Statistical association*, 71, 210-213.
10. Kusum, K. (1985). A new distribution-free test for two-sample scale problem. *J. Indian Statistical Assoc*. 23: 97-107.
11. Moses, L. E. (1963). Rank Tests of Dispersion. *The Annals of Mathematical Statistics*. 34, 973-983.
12. Nelson, P. R. (1982). Exact critical points for the analysis of means. *Communications in Statistics-Theory and Methods*. 11(6), 699-709.
13. Nelson, P. R. (1993). Additional uses for the analysis of means and extended tables of critical values. *Technometrics*, 35, 61-71.
14. Nelson, P. N., Wludyka, P. S and Copeland, K. A. F. (2005). *Analysis of means: a graphical method for comparing means, rates, and proportions*. SIAM Press, PA.
15. Ott, E. R. (1967). Analysis of Means – A Graphical Procedure. *Industrial Quality Control*, 24, 101-109.
16. Randles, R. H. and Wolfe, D. A. (1997). *Introduction to the Theory of Nonparametric Statistics*. John Wiley & Sons, New York, NY.
17. Shetty, I. D. and Pandit, P. V. (2004). A Note on a new class of distribution-free tests for the two-sample scale problem based on subsample medians. *Communications in Statistics-Theory and Methods*. 33(9), 2271-2280.
18. Wludyka, P. S. and Nelson, P. N. (1999). Two non-parametric, analysis–of-means-type tests for homogeneity of variances. *Journal of Applied Statistics*. 26(2), 243-256.